Chemical Data Assimilation with CMAQ: Continuous vs. Discrete Advection Adjoints

Tianyi Gou, Kumaresh Singh, and Adrian Sandu

Department of Computer Science, Virginia Tech.,
Blacksburg, VA, 24060, USA
{tygou,kumaresh,sandu}@cs.vt.edu

Abstract. The Community Multiscale Air Quality (CMAQ) system is the Environmental Protection Agency’s main modeling tool for atmospheric pollution studies. CMAQ-ADJ, the adjoint model of CMAQ, offers new capabilities such as receptor-oriented sensitivity analysis and chemical data assimilation. This paper presents the construction of discrete advection adjoints in CMAQ. The new adjoints are thoroughly validated against finite differences. We assess the performance of discrete and continuous advection adjoints in CMAQ on sensitivity analysis and 4D-Var data assimilation applications. The results show that discrete adjoint sensitivities better agree with finite difference value than their continuous counterparts. However, continuous adjoints result in a faster convergence of the numerical optimization in 4D-Var data assimilation. Similar conclusions apply to modified discrete adjoints.

Keywords: Chemical Transport Models, Adjoints, Data Assimilation.

1 Introduction

The Community Multiscale Air Quality (CMAQ) modeling system is a powerful third generation air quality modeling tool used to predict the state of the air-borne pollutants [1]. CMAQ is capable of modeling important air quality issues such as tropospheric ozone, fine particles, acid deposition, and visibility degradation [2]. CMAQ is the Environmental Protection Agency’s main modeling tool for assessing atmospheric pollution, and it is being widely used by the scientific community for a variety of air quality studies.

Adjoint modeling is a powerful tool for computing the sensitivities of the model output with respect to (thereafter, w.r.t.) a large number of model inputs. Adjoints are widely used in applications including sensitivity analysis, data assimilation, parameter estimation, stability analysis, etc. [6]. There are two approaches to adjoint model development. In the *continuous adjoint* approach one first differentiates the underlying mathematical equations at an abstract level, then discretizes the resulting adjoint equations. In the *discrete adjoint* approach one first discretizes the physical equations (using a suitable numerical method), then differentiates the discrete algorithm. Discrete adjoints are popular since they can be obtained by automatic differentiation [7]. The two approaches lead to different adjoint models.

CMAQ-ADJ is the adjoint model of the CMAQ modeling system, and has been developed through a collaboration between Caltech, Virginia Tech, and University of Houston [2,9,10]. The CMAQ adjoint system has the added capability of performing sensitivity analysis and 4D-Var data assimilation. The current implementation of CMAQ-ADJ considers the advection, diffusion, emission, deposition, and gas-phase processes. The continuous adjoint approach is used for advection processes, while the discrete adjoint approach is used for other processes including chemistry process and diffusion process.

In this paper we present the construction and validation of two newly developed modules CMAQ-ADJ: one is a discrete advection adjoint process, and the second is a modified discrete adjoint, where we remove the monotonic property of the piecewise parabolic method (PPM) [8] used to discretize the horizontal advection process. For the remainder of this paper, we refer to the first one as discrete adjoint version, to the latter one as modified discrete adjoint version, and to the current CMAQ-ADJ advection as the continuous adjoint version. Differences of three adjoint versions for sensitivity analysis and 4D-Var data assimilation are illustrated as well.

This paper is organized as follows. Section 2 introduces the chemical transport model CMAQ. Section 3 presents the construction of discrete and continuous adjoints, as well as their validation. Applications to sensitivity analysis and 4D-Var data assimilation are presented in sections 4 and 5, respectively. Section 6 provides a summary of the work presented in this paper.

2 Chemical Transport Model in CMAQ

Atmospheric Chemical Transport Models (CTMs) are used to capture the knowledge about the physical and chemical processes that govern the evolution of airborne pollutant fields. They have the ability to predict the concentrations of the airborne pollutants. Atmospheric CTMs solve the atmospheric mass balance equation defined as follows [2]:

$$\frac{\partial C_i}{\partial t} = -\mathbf{u} \cdot \nabla C_i + \frac{1}{\rho} \nabla \cdot (\rho \mathbf{K} \nabla C_i) + R_i + E_i \quad (1)$$

where C_i is the concentration for species i , $\mathbf{u}$ is the wind fields, ρ is the air density, $\mathbf{K}$ is the diffusivity tensor, R_i represents the chemical reaction rate for species i and E_i represents emissions. From the above formula, we see that CTMs model the following important processes: advection ($-\mathbf{u} \cdot \nabla C_i$), diffusion ($\rho^{-1} \nabla \cdot (\rho \mathbf{K} \nabla C_i)$), chemistry (R_i) and emission (E_i). In addition to these four parts, CMAQ also includes cloud process, plume-in-grid process and aerosol process. Details of all the science processes can be found in [1,5].

To develop the continuous adjoint system of a CTM we first derive its tangent linear model (TLM) by differentiation, then use Lagrange multipliers and integration by parts to get the adjoint system. Details of deriving adjoint system can be found in [3].

3 Adjoint Construction and Validation

The adjoint of a multi-physics forward CTM requires the construction of adjoints for each of its science processes. As mentioned in the first section, two approaches can be used to develop adjoint model: continuous adjoint approach and discrete adjoint approach.

In this paper we report on a new implementation of a discrete advection adjoint process. Details on adjoint development for other processes can be found in [2].

3.1 Vertical Advection

The vertical advection equation in CMAQ [2] is

$$\frac{\partial(\rho C)}{\partial t} = -\frac{\partial(w\rho C)}{\partial z} \quad (2)$$

where C is concentration vector (a function of time and space), ρ is the air density and w is the vertical wind field.

Continuous Adjoint of Vertical Advection. The corresponding continuous adjoint model of vertical advection [2] is

$$-\frac{\partial(\lambda/\rho)}{\partial t} = \frac{\partial(w\lambda/\rho)}{\partial z} \quad (3)$$

where λ is the adjoint variable.

Discrete Adjoint of Vertical Advection. Discrete adjoint of vertical advection is implemented from the forward code with the aid of automatic differential tool such as TAMC (Tangent linear and Adjoint Model Compiler) [7]. For the adjoint mode, we need checkpoints of the forward concentration state. Therefore, we write the concentration into checkpoint files after every dynamic time step in forward mode and read them in the adjoint mode before every dynamic time step.

Validation of Vertical Advection Adjoints. To validate the adjoints of vertical advection, we compare the adjoint sensitivity with central finite difference sensitivity. Specifically, we choose a source layer ($L_s = 7$), a receptor layer ($L_r = 10$), and a species ($S = O_3$, ozone). A perturbation of species S is introduced in each grid of the source layer and is tracked at the corresponding grid (i.e., in the same column) at the receptor layer. In other words, we compute $dC(L_r, S)/dC(L_s, S)$ (for the same column) using both central finite difference approach and the adjoint sensitivity approach. Fig. 1(a) shows the scattered plot of continuous adjoint sensitivity vs. central finite difference values for 8 hours run of CMAQ-ADJ with vertical advection only. Each point represents one sensitivity (one column), with the x coordinate the value obtained by the adjoint method and the y coordinate the value obtained by central differences. Fig. 1(b) shows a similar scattered plot for discrete adjoint sensitivities. We see that discrete adjoint sensitivities of vertical advection agree better with finite difference results than continuous adjoint sensitivities.

Fig. 1. Scattered plot of (a) continuous adjoint sensitivity and (b) discrete adjoint sensitivity vs. finite difference values for an 8 hours run with vertical advection only. Source layer $L_s = 7$, receptor layer $L_r = 10$, and species $S = O_3$ (ozone). Each point represents one sensitivity (one column), with the x coordinate the value obtained by the adjoint method and the y coordinate the value obtained by central differences.

3.2 Horizontal Advection

In CMAQ horizontal advection is directionally split, with successive calls made to x-axis advection and to y-axis advection. Since the same one-dimensional advection algorithm is used for both x and y directions, we consider x-advection in this section. The one-dimensional horizontal advection equation [2] is similar to (2), with the horizontal wind field u replacing w . If the total mass continuity holds for the horizontal advection equation, its equivalent form is

$$\frac{\partial(C)}{\partial t} = -u \frac{\partial(C)}{\partial x} \quad (4)$$

Continuous Adjoint of Horizontal Advection. The corresponding continuous adjoint of horizontal advection takes the form (3) with λ the adjoint variable, ρ the air density, and u (instead of w) the horizontal wind field. The continuous adjoint equation is implemented by calling the forward horizontal advection subroutine with a reversed wind field ($-u$).

Discrete Adjoint of Horizontal Advection. The same automatic-differentiation based approach used to develop the discrete adjoint of vertical advection is employed to develop the discrete adjoint of horizontal advection.

Modified Discrete Adjoint of Horizontal Advection. In CMAQ the numerical method used to solve the advection equation is the piecewise parabolic method (PPM). PPM has monotonicity properties obtained with the help of slope and curvature limiters. The reverse mode differentiation of these limiters may introduce discontinuities in the discrete adjoint solution [8].

In the modified discrete adjoint approach we remove monotonicity characteristics of the PPM by commenting out the steepening procedure (i.e., the limiters);

this non-monotonic code is then processed by automatic differentiation. The algorithm is smooth and has no points of non-differentiability.

Validation of Horizontal Advection Adjoints. To validate adjoint of horizontal advection, we plot the adjoint sensitivities and finite difference values. Specifically, we choose a source column (c_s), a receptor column (c_r), and a species ($S = O_3$). For a given vertical layer and all horizontal rows, the sensitivities $dC(c_r, S)/dC(c_s, S)$ are computed using adjoints and using central finite differences.

Fig. 2 presents the scattered plots of all three adjoint sensitivities against the finite difference counterparts. An 8-hour horizontal advection experiment is carried out. Each point represents a sensitivity pair at different vertical layer and horizontal row. The continuous adjoint results are shown in Fig. 2 (a), the discrete adjoints in Fig. 2 (b), and the modified discrete adjoints in Fig. 2 (c) shows the scatter plot of the modified discrete adjoint vs. finite difference results. The continuous adjoint sensitivities show the highest discrepancies against the finite differences. Both the discrete adjoint and the modified discrete adjoint results show good agreement with finite difference results.

(a) Continuous adj. vs. Finite diff.

(b) Discrete adj. vs. Finite diff.

(c) Modified discrete adj. vs. Finite diff.

Fig. 2. Scattered plot of (a) continuous adjoint sensitivity, (b) discrete adjoint sensitivity, and (c) modified discrete adjoint sensitivity vs. finite difference values for 8 hours run of x-horizontal advection only with source column 25, receptor column 23 and species ozone

4 Sensitivity Analysis

Sensitivity analysis quantifies changes in the model output w.r.t. the variations in model inputs. The sensitivity (derivative) of the model output/receptor with respect to model inputs/sources can be obtained by two approaches. One is the forward approach where the perturbation of the input/source propagates forward in time to the output/receptor. The other one is adjoint approach where the perturbation of the receptor is traced backward to the contributing sources. Therefore the forward approach can efficiently compute the sensitivity of many outputs/receptors w.r.t. a few sources. The adjoint approach can efficiently provide the sensitivities of a few outputs/receptors w.r.t. many inputs/sources.

Fig. 3. Sensitivity of O_3 at localized receptors at July 3, 1999, 00:00:00 w.r.t. NO_2 source on July 2, 1999, 18:00:00. O_3 receptor location is cols = 18:22, rows = 18:22, layer = 1:4.

With the development of CMAQ-ADJ, CMAQ is capable of performing adjoint sensitivity analysis. To see the difference of sensitivities provided by the three different versions of advection adjoint Fig. 3 presents the results of a 24 hours simulation. The model output is the sum of O_3 concentrations in the area described by grid cells (18 : 22, 18 : 22, 1 : 4), at 00:00:00 (GMT) hours on July 3, 1999. The model inputs are initial NO_2 conditions at 18:00:00 (GMT) hours on July 2, 1999. Fig. 3 reveals small differences between the sensitivity fields computed by the three adjoint versions. We next study how this difference affects the 4D-Var application.

5 4D-Var Data Assimilation

Data assimilation is the process of fusing information from model predictions and observations in order to obtain better initial conditions, boundary conditions or emission estimates. 4D-Var data assimilation optimally combines three sources of information: a priori estimate of the state of the atmosphere; knowledge about the physical and chemical processes that govern the evolution of pollutant fields as captured in the model; and observations of some of the state variables [3].

We apply 4D-Var data assimilation with CMAQ-ADJ to provide optimal estimates of the initial conditions. 4D-Var data assimilation is posed as an optimization problem where the best estimate of initial conditions minimizes the following cost function:

$$J(c^0) = \frac{1}{2} \sum_{k=1}^N (c^k - c^{k,obs})^T R_k^{-1} (c^k - c^{k,obs}) + \frac{1}{2} (c^0 - c^b)^T B^{-1} (c^0 - c^b) \quad (5)$$

where c^0 is the initial concentration, c^k is the model prediction at time step k , $c^{(k,obs)}$ is the observation at time step k , R_k is the observation error covariance matrix, c^b is the background concentration and B is the background error covariance matrix. The cost function measures the misfit between model predictions and observations as well as the misfit between initial conditions and background concentrations.

5.1 Implementation

To minimize the cost function (5), we apply L-BFGS, a limited memory quasi-Newton method for solving large scale optimization problems [4]. L-BFGS is an iterative method which requires, at each iteration step, the cost function value and its gradient at that point, and returns another point which is a better approximation to the optimal solution. The whole process continues until the convergence criteria are satisfied.

In the implementation of CMAQ/4D-Var we interface L-BFGS with the CMAQ-ADJ. The cost function value and gradient value are obtained through a forward run followed by an adjoint run of CMAQ-ADJ.

Fig. 4. 4D-Var data assimilation carried out for 12 hours to optimize the initial O_3 values. (a) difference between perturbed and reference concentration at initial time; (b) - (d) difference between optimized and reference concentration at initial time using (b) continuous adjoint; (c) discrete adjoint; (d) modified discrete adjoint.

5.2 4D-Var Results

To validate our CMAQ/4D-Var implementation we have designed a test case as follows. First, we choose a reference initial concentration c_0 and perform a forward model run to lay down a set of synthetic observations. Second, a perturbation is introduced to obtain $c_p = c_0 + \Delta c_0$. Third, $c^b = c_p$ is used as the background value and as the starting point for L-BFGS algorithm. After each iteration run, we obtain a new initial point c_0^k . Again, by performing a forward run and an adjoint run, we obtain the cost function value and gradient at c_0^k , which are used as the input for the next L-BFGS iteration. When convergence occurs we get an optimal estimate c_0^a to the reference initial condition c_0 .

Fig. 5. Convergence of L-BFGS optimization. (a) Log of the cost function values vs. model runs; (b) RMS values vs. model runs.

In the following section, we show the 4D-Var simulation results for three different versions of CMAQ-ADJ. In our experiment, we perturbed ozone and retrieved ozone only. We visualize the difference between the initial guess and the reference as well as the difference between the optimized solution and the reference. In addition, we plot the log of the cost function values as the iteration number increases. To see if the optimized solution converges to the reference, we also plot the root-mean square (RMS) values given by: $\text{RMS} = \|c_0^k - c_0\| / \|c_0\|$.

Fig.4 and Fig.5 show the 4D-Var validation results for a 12 hours simulation with all science processes for species O_3 . With the continuous adjoint the optimized initial condition agrees very well with the reference initial condition. In addition, the cost function and RMS values decrease significantly with the number of model runs indicating a fast convergence to the reference initial condition. On the other hand, the discrete adjoint shows slow convergence, and the modified discrete adjoint does not achieve convergence.

6 Conclusion

In this paper we present the construction and validation of discrete and modified discrete advection adjoints in CMAQ. Discrete adjoint sensitivities given by both approaches match the finite difference results better than continuous adjoint sensitivities. However, in 4D-Var data assimilation experiments, a much faster convergence is obtained when the gradients are computed using the continuous adjoint approach than using the discrete adjoint or modified adjoint approaches. We believe that this effect is related to the smoothness of the respective gradients. Accordingly, we conclude that, for CMAQ-ADJ, continuous adjoint is preferable for 4D-Var data assimilation while discrete adjoint provide better sensitivity agreement with finite difference results. Both continuous and discrete adjoint implementations are available to the user in the new CMAQ-ADJ version.

References

1. Byun, D.W., Ching, J.K.S.: Science Algorithms of the EPANet-3 Community Multiscale Air Quality (CMAQ) Modeling System; U.S. EPA/600/R-99/030; U.S. Environmental Protection, Agency: Research Triangle Park, NC (1999)
2. Hakami, A., Henze, D.K., Seinfeld, J.H., Singh, K., Sandu, A., Kim, S., Byun, D., Li, Q.: The Adjoint of CMAQ. *Environ. Sci. Technol.* 41(22), 7807–7817 (2007)
3. Sandu, A., Daescu, D.N., Carmichael, G.R., Chai, T.F.: Adjoint sensitivity analysis of regional air quality models. *J. Comput. Phys.* 204, 222–252 (2005)
4. Liu, D.C., Nocedal, J.: On the limited memory BFGS method for large-scale optimization. *Math. Programming* 45, 503–528 (1989)
5. Byun, D.W., Schere, K.: Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. *Applied Mechanics Reviews* 59, 51–77 (2006)
6. Errico, R.M.: What is an adjoint model? *Bull. Amer. Meteor. Soc.* 78, 2577–2591 (1997)
7. Giering, R., Kaminski, T.: Recipes for Adjoint code Construction. *ACM Trans. Math. Softw.* 24, 437–474 (1998)
8. Colella, P., Woodward, P.R.: The Piecewise Parabolic Method (PPM) for Gas-Dynamical Simulations. *J. Comput. Phys.* 54, 174–201 (1984)
9. Singh, K., Sandu, A., Hakami, A. and Seinfeld: CMAQ-ADJv4.5 (2007), http://people.cs.vt.edu/~asandu/Software/CMAQ_ADJ/CMAQ_ADJ.html
10. Singh, K., Sandu, A., Hakami, A., Seinfeld, J.: CMAQ v4.5 Adjoint Users's Manual (2007)